Solutions for the MIMO Gaussian Wiretap Channel with a Cooperative Jammer

S. Ali. A. Fakoorian*, *Student Member, IEEE* and A. Lee Swindlehurst, *Fellow, IEEE*

Abstract

We study the Gaussian MIMO wiretap channel with a transmitter, a legitimate receiver, an eavesdropper and an external helper, each equipped with multiple antennas. The transmitter sends confidential messages to its intended receiver, while the helper transmits jamming signals independent of the source message to confuse the eavesdropper. The jamming signal is assumed to be treated as noise at both the intended receiver and the eavesdropper. We obtain a closed-form expression for the structure of the artificial noise covariance matrix that guarantees no decrease in the secrecy capacity of the wiretap channel. We also describe how to find specific realizations of this covariance matrix expression that provide good secrecy rate performance, even when there is no non-trivial null space between the helper and the intended receiver. Unlike prior work, our approach considers the general MIMO case, and is not restricted to SISO or MISO scenarios.

Index Terms

Physical-layer security, interference channel, MIMO wiretap channel, cooperative jamming.

EDICS: WIN-CONT, WIN-PHYL, WIN-INFO, MSP-CAPC

The authors are with the Dept. of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697-2625, USA. e-mail:{afakoori, swindle}@uci.edu

This work was supported by the U.S. Army Research Office under the Multi-University Research Initiative (MURI) grant W911NF-07-1-0318.

I. Introduction

Recent information-theoretic research on secure communication has focused on enhancing security at the physical layer. The wiretap channel, first introduced and studied by Wyner [1], is the most basic physical layer model that captures the problem of communication security. This work led to the development of the notion of perfect secrecy capacity, which quantifies the maximum rate at which a transmitter can reliably send a secret message to its intended recipient, without it being decoded by an eavesdropper. The Gaussian wiretap channel, in which the outputs of the legitimate receiver and the eavesdropper are corrupted by additive white Gaussian noise, was studied in [2]. The secrecy capacity of a Gaussian wiretap channel, which is in general a difficult non-convex optimization problem, has been addressed and solved for in [3]-[7]. The secrecy capacity under an average power constraint is treated in [4] and [5], where in [4] a beamforming approach, based on the generalized singular value decomposition (GSVD), is proposed that achieves the secrecy capacity in the high SNR regime. In [5], we propose an optimal power allocation that achieves the secrecy capacity of the GSVD-based multiple-input, multiple-output (MIMO) Gaussian wiretap channel for any SNR. In [7], a closed-form expression for the secrecy capacity is derived under a certain power-covariance constraint.

It was shown in [8] that, for a wiretap channel without feedback, a non-zero secrecy capacity can only be obtained if the eavesdropper's channel is of lower quality than that of the intended recipient. Otherwise, it is infeasible to establish a secure link under Wyner's wiretap channel model. In such situations, one approach is to exploit user cooperation in facilitating the transmission of confidential messages from the source to the destination. In [9]-[13], for example, a four-terminal relay-eavesdropper channel is considered, where a source wishes to send messages to a destination while leveraging the help of a relay/helper node to hide the messages from the eavesdropper. While the relay can assist in the transmission of confidential messages, its computational cost may be prohibitive and there are difficulties associated with the coding and decoding schemes at both the relay and the intended receiver. Alternatively, a cooperating node can be used as a helper that simply transmits jamming signals, independent of the source message, to confuse the eavesdropper and increase the range of channel conditions under which secure communications can take place. The strategy of using a helper to improve the secrecy of the source-destination communication is generally known as cooperative jamming [9], [11] or noise-forwarding [12] in prior work.

In [9], the scenario where multiple single-antenna users communicate with a common receiver (i.e., the multiple access channel) in the presence of an eavesdropper is considered, and the optimal transmit

power allocation that achieves the maximum secrecy sum-rate ia obtained. The work of [9] shows that any user prevented from transmitting based on the obtained power allocation can help increase the secrecy rate for other users by transmitting artificial noise to the eavesdropper (cooperative jamming). In [11], a source-destination system in the presence of multiple helpers and multiple eavesdroppers is considered, where the helpers can transmit weighted jamming signals to degrade the eavesdropper's ability to decode the source. While the objective is to select the weights so as to maximize the secrecy rate under a total power constraint, or to minimize the total power under a secrecy rate constraint, the results in [11] yield sub-optimal weights for both single and multiple eavesdroppers, due to the assumption that the jamming signal must be nulled at the destination. The noise forwarding scheme of [12] requires that the interferer's codewords be decoded by the intended receiver. A generalization of [9], [11] and [12] is proposed in [13], in which the helper's codewords do not have to be decoded by the receiver.

The prior work in [9]-[13] assumes single antenna nodes and models single-input, single-output (SISO) or multiple-input, single-output (MISO) cases. A more general MIMO case with multiple cooperative jammers was studied in [14], in which the jammers aligned their interference to lie within a pre-specified "jamming subspace" at the receiver, but the dimensions of the subspace and the power allocation were not optimized. In this paper, we also address the general MIMO case, where the transmitter, legitimate receiver, eavesdropper and helper are in general all equipped with multiple antennas. The transmitter sends confidential messages to its intended receiver, while the helper node assists the transmitter by sending jamming signals independent of the source message to confuse the eavesdropper. While the previous work on this problem shows the fundamental role of jamming as a means to increase secrecy rates, it also emphasizes the fact that that non-carefully designed jamming strategies can preclude secure communication [15].

In this work, we derive a closed-form expression for the structure of the artificial noise covariance matrix of a cooperating jammer that guarantees no decrease in the secrecy capacity of the wiretap channel, assuming the jamming signal from the helper is treated as noise at both the intended receiver and the eavesdropper. We describe algorithms for finding specific realizations of this covariance expression that provide good secrecy rate performance, and show that even when there is no non-trivial nullspace between the helper and the intended receiver, the helper can still transmit artificial noise that does not impact the mutual information between the transmitter and the intended receiver, while decreasing the mutual information between the transmitter and the eavesdropper. Hence, the secrecy level of the confidential message is increased. The situation we consider is different from the one in [16], where the transmitter itself rather than an external helper broadcasts artificial noise to degrade the eavesdropper's channel.

However, both approaches are able to achieve a positive perfect secrecy rate in scenarios where the secrecy capacity in the absence of jamming is zero.

The remainder of the paper is organized as follows. In Section II, we describe the system model for the helper-assisted Gaussian MIMO wiretap channel and formulate the problem to be solved. In Sections III and IV, we derive the artificial noise covariance matrix that guarantees no decrease in the secrecy capacity of the wiretap channel. Numerical results in Section V are presented to illustrate the proposed solution. Finally, Section VI concludes the paper.

Notation: Throughout the paper, we use boldface uppercase letters to denote matrices. Vector-valued random variables are written with non-boldface uppercase letters (*e.g.,* X), while the corresponding lowercase boldface letter ($\mathbf{x}$) denotes a specific realization of the random variable. Scalar variables are written with non-boldface (lowercase or uppercase) letters. We use $(.)^T$ to represent matrix transposition, $(.)^H$ the Hermitian (i.e., conjugate) transpose, Tr(.) the matrix trace, E the expectation operator, $\mathbf{I}$ the identity matrix, and $\mathbf{0}$ a matrix or vector with all zeros. Mutual information between the random variables A and B is denoted by $I(A;B)$, and $\mathcal{CN}(0,1)$ represents the complex circularly symmetric Gaussian distribution with zero mean and unit variance.

II. System Model

We consider a MIMO wiretap channel that includes a transmitter, an intended receiver, a helping interferer and an eavesdropper, with n_t, n_r, n_h and n_e antennas, respectively. The transmitter sends a confidential message to the intended receiver with the aid of the helper, in the presence of an eavesdropper. We assume that the helper does not know the confidential message and transmits only a Gaussian jamming signal which is not known at the intended receiver nor the eavesdropper and which is treated as noise at both receivers. The mathematical model for this scenario is given by:

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x}_1 + \mathbf{G}_2\mathbf{x}_2 + \mathbf{z}_1 \quad (1)$$

$$\mathbf{y}_2 = \mathbf{H}_2\mathbf{x}_2 + \mathbf{G}_1\mathbf{x}_1 + \mathbf{z}_2\ , \quad (2)$$

where $\mathbf{x}_1$ is a zero-mean $n_t \times 1$ transmitted signal vector, $\mathbf{x}_2$ is a zero-mean $n_h \times 1$ jamming vector transmitted by the helper, and $\mathbf{z}_1 \in \mathbb{C}^{n_r \times 1}$, $\mathbf{z}_2 \in \mathbb{C}^{n_e \times 1}$ are additive white Gaussian noise (AWGN) vectors at the intended receiver and the eavesdropper, respectively, with i.i.d. entries distributed as $\mathcal{CN}(0,1)$. The matrices $\mathbf{H}_1, \mathbf{G}_1$ represent the channels from the transmitter to the intended receiver and eavesdropper, respectively, while $\mathbf{H}_2, \mathbf{G}_2$ are the channels from the helper to the eavesdropper and intended receiver, respectively. The channels are assumed to be independent of each other and full rank with arbitrary

dimensions. We also assume that the transmitter has full channel state information and is aware of the effective noise covariance at both receivers, where the effective noise is the background noise plus the received artificial noise. Both the helper and the eavesdropper are also aware of all channel matrices as well.

The jamming signal transmitted by the helper satisfies an average power constraint:

$$\text{Tr}(E\{X_2 X_2^H\}) = \text{Tr}(\mathbf{K}_w) \leq P_h \tag{3}$$

where X_2 is the random variable associated with the specific realization $\mathbf{x}_2$ and $\mathbf{K}_w$ is the corresponding covariance matrix. The channel input is subject to a matrix power constraint [7], [17]

$$E\{X_1 X_1^H\} = \mathbf{K}_x \preceq \mathbf{S} \tag{4}$$

where $\mathbf{K}_x$ is the input covariance matrix, $\mathbf{S}$ is a positive semi-definite matrix, and "$\preceq$" denotes that $\mathbf{S} - \mathbf{K}_x$ is positive semi-definite. Note that (4) is a rather general power constraint that subsumes many other important power constraints, including the average total and per-antenna power constraints as special cases. The approach developed in this paper will assume that P_h and $\mathbf{S}$ (or $\text{Tr}(\mathbf{S}) \leq P_t$) are fixed, and that power is not allocated jointly between the transmitter and helper. The numerical results presented later, however, will illustrate the trade-off associated with the power allocation when $P_h + P_t$ is fixed.

As mentioned before, we assume Gauusian signaling for the helper. Thus the effective noise at both receivers is Gaussian and consequently the above MIMO wiretap channel model is Gaussian. For this case, a Gaussian input signal is the optimal choice [6], [17]. Hence, the general optimization problem is equivalent to finding the matrices $\mathbf{K}_x \succeq 0$ and $\mathbf{K}_w \succeq 0$ that allow the secrecy capacity of the network to be obtained. A matrix characterization of this optimization problem is given by:

$$\begin{aligned} C_{sec} &= \max_{\mathbf{K}_x \succeq 0, \mathbf{K}_w \succeq 0} [I(X_1; Y_1) - I(X_1; Y_2)] \\ &= \max_{\mathbf{K}_x \succeq 0, \mathbf{K}_w \succeq 0} \log |\mathbf{K}_x \mathbf{H}_1^H (\mathbf{G}_2 \mathbf{K}_w \mathbf{G}_2^H + \mathbf{I})^{-1} \mathbf{H}_1 + \mathbf{I}| \\ &\qquad - \log |\mathbf{K}_x \mathbf{G}_1^H (\mathbf{H}_2 \mathbf{K}_w \mathbf{H}_2^H + \mathbf{I})^{-1} \mathbf{G}_1 + \mathbf{I}| \,, \end{aligned} \tag{5}$$

where the non-convex maximization problem in carried out under the power constraints given in (3) and (4).

Lemma 1: For a given $\mathbf{K}_w$, the maximum of (5) is given by

$$C_{sec}(\mathbf{S}) = \sum_{i=1}^{\rho} \log \gamma_i \tag{6}$$

where γ_i, $i = 1, \cdots, \rho$, are the generalized eigenvalues of the pencil

$$(\mathbf{S}^{\frac{1}{2}} \mathbf{H}_1^H (\mathbf{G}_2 \mathbf{K}_w \mathbf{G}_2^H + \mathbf{I})^{-1} \mathbf{H}_1 \mathbf{S}^{\frac{1}{2}} + \mathbf{I}, \quad \mathbf{S}^{\frac{1}{2}} \mathbf{G}_1^H (\mathbf{H}_2 \mathbf{K}_w \mathbf{H}_2^H + \mathbf{I})^{-1} \mathbf{G}_1 \mathbf{S}^{\frac{1}{2}} + \mathbf{I}) \tag{7}$$

that are greater than 1.

Proof: When the optimization problem in (5) is performed over $\mathbf{K}_x$ under the matrix power constraint (4) for a given $\mathbf{K}_w$, it is equivalent to a simple MIMO Gaussian wiretap channel without a helper, where the noise covariance matrices at the receiver and the eavesdropper are $(\mathbf{G}_2\mathbf{K}_w\mathbf{G}_2^H+\mathbf{I})$ and $(\mathbf{H}_2\mathbf{K}_w\mathbf{H}_2^H+\mathbf{I})$, respectively. The above lemma is a natural extension of [7] and [17, Theorem 3] for the standard MIMO Gaussian wiretap channel.

Note that since both elements of the pencil (7) are strictly positive definite, all of the generalized eigenvalues are real and positive [17], [18]. In (6), a total of ρ of them are assumed to be greater than one. Clearly, if there are no such eigenvalues, then the information signal received at the intended receiver is a degraded version of that of the eavesdropper, and in this case the secrecy capacity is zero. Note also that Lemma 1 only provides the secrecy capacity for the optimal $\mathbf{K}_x$, but does not give an explicit expression for this $\mathbf{K}_x$. A general expression for the maximizing $\mathbf{K}_x$ will be given in the next section.

To solve the general optimization problem in (5), we would need to find the $\mathbf{K}_w$ that maximizes (6). Unfortunately, this appears to be a very difficult problem to solve without resorting to some type of *ad hoc* search. In the following we obtain a sub-optimal closed-form solution for the artificial noise covariance matrix $\mathbf{K}_w$ that guarantees no decrease in the mutual information between the transmitter and the intended receiver compared with the case where $\mathbf{K}_w = \mathbf{0}$, while maintaining the power constraint in (5). Hence, the new non-zero $\mathbf{K}_w$ will only interfere with the eavesdropper, and the secrecy level of the confidential message will be increased. Once such a $\mathbf{K}_w$ is found, additional improvement in the secrecy rate can be achieved if the transmitter updates its covariance matrix $\mathbf{K}_x$ for the obtained $\mathbf{K}_w$. The final secrecy rate for this method is obtained by simply computing (6) and (7) for the resulting $\mathbf{K}_w$. Note that we will not propose an iterative algorithm that would further alternate between calculating $\mathbf{K}_x$ and $\mathbf{K}_w$. We will see in the next section that there is no clear way to update $\mathbf{K}_w$ from a known non-zero value.

III. Analytical Method

We begin with the case where the helper transmits no signal $(\mathbf{K}_w = 0)$. In this case, the communication system is reduced to a simple MIMO Gaussian wiretap channel without helper. Based on Lemma 1, the maximum of (5) when $\mathbf{K}_w = 0$ is obtained by applying the generalized eigenvalue decomposition to the following two Hermitian positive definite matrices [7], [17]:

$$\mathbf{S}^{\frac{1}{2}}\mathbf{H}_1^H\mathbf{H}_1\mathbf{S}^{\frac{1}{2}}+\mathbf{I}, \quad \mathbf{S}^{\frac{1}{2}}\mathbf{G}_1^H\mathbf{G}_1\mathbf{S}^{\frac{1}{2}}+\mathbf{I}\,.$$

In particular, there exists an invertible generalized eigenvector matrix $\mathbf{C}$ such that [18]

$$\mathbf{C}^H \left[\mathbf{S}^{\frac{1}{2}}\mathbf{G}_1^H\mathbf{G}_1\mathbf{S}^{\frac{1}{2}} + \mathbf{I}\right]\mathbf{C} = \mathbf{I} \tag{8}$$

$$\mathbf{C}^H \left[\mathbf{S}^{\frac{1}{2}}\mathbf{H}_1^H\mathbf{H}_1\mathbf{S}^{\frac{1}{2}} + \mathbf{I}\right]\mathbf{C} = \mathbf{\Lambda} \tag{9}$$

where $\mathbf{\Lambda} = \text{diag}\{\lambda_1, ..., \lambda_{n_t}\}$ is a positive definite diagonal matrix and $\lambda_1, ..., \lambda_{n_t}$ represent the generalized eigenvalues. Without loss of generality, we assume the generalized eigenvalues are ordered as

$$\lambda_1 \geq ... \geq \lambda_b > 1 \geq \lambda_{b+1} \geq ... \geq \lambda_{n_t} > 0$$

so that a total of b $(0 \leq b \leq n_t)$ are assumed to be greater than 1. Hence, we can write $\mathbf{\Lambda}$ as

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \tag{10}$$

where $\mathbf{\Lambda}_1 = \text{diag}\{\lambda_1, ..., \lambda_b\}$ and $\mathbf{\Lambda}_2 = \text{diag}\{\lambda_{b+1}, ..., \lambda_{n_t}\}$. Also, we can write $\mathbf{C}$ as

$$\mathbf{C} = [\mathbf{C}_1 \quad \mathbf{C}_2] \tag{11}$$

where $\mathbf{C}_1$ is the $n_t \times b$ submatrix representing the generalized eigenvectors corresponding to $\{\lambda_1, ..., \lambda_b\}$ and $\mathbf{C}_2$ is the $n_t \times (n_t - b)$ submatrix representing the generalized eigenvectors corresponding to $\{\lambda_{b+1}, ..., \lambda_{n_t}\}$.

For the case of $\mathbf{K}_w = 0$, the secrecy capacity of (5) under the matrix power constraint (4) is given by (Lemma 1 or [17, Theorem 3]):

$$C_{sec} = \sum_{i=1}^{b} \log \lambda_i = \log |\mathbf{\Lambda}_1| \tag{12}$$

and the input covariance matrix $\mathbf{K}_x^*$ that maximizes (5) is given by ([7], [17]):

$$\mathbf{K}_x^* = \mathbf{S}^{\frac{1}{2}}\mathbf{C} \begin{bmatrix} (\mathbf{C}_1^H\mathbf{C}_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{C}^H\mathbf{S}^{\frac{1}{2}} \ . \tag{13}$$

Note that (13) is a general expression for the $\mathbf{K}_x$ that optimizes (5) for a given $\mathbf{K}_w$ even when $\mathbf{K}_w \neq 0$, although in this case the $\mathbf{C}$ will be the generalized eigenvector matrix of the pencil (7). From (9) we note that $\mathbf{H}_1^H\mathbf{H}_1$ can be written as

$$\mathbf{H}_1^H\mathbf{H}_1 = \mathbf{S}^{-1/2}\left[\mathbf{C}^{-H}\begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix}\mathbf{C}^{-1} - \mathbf{I}\right]\mathbf{S}^{-1/2} \ . \tag{14}$$

The following lemma gives the mutual information $I(X_1; Y_1)$ between the transmitter and the intended receiver when $\mathbf{K}_w = 0$ and $\mathbf{K}_x$ is given by (13).

Lemma 2: The following equality holds:

$$I(X_1; Y_1)|_{\mathbf{K}_w=0, \mathbf{K}_x=\mathbf{K}_x^*} = \log\left|\mathbf{K}_x^*\mathbf{H}_1^H\mathbf{H}_1 + \mathbf{I}\right| = \log\left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1\right| \ . \tag{15}$$

Proof: Following the same steps as the proof of [7, App. D] and using (13) and (14), we have

$$\begin{aligned}
\left|\mathbf{K}_x^*\mathbf{H}_1^H\mathbf{H}_1+\mathbf{I}\right| &= \left|\mathbf{S}^{\frac{1}{2}}\mathbf{C}\begin{bmatrix}(\mathbf{C}_1^H\mathbf{C}_1)^{-1} & 0\\ 0 & 0\end{bmatrix}\mathbf{C}^H\times\left[\mathbf{C}^{-H}\begin{bmatrix}\mathbf{\Lambda}_1 & 0\\ 0 & \mathbf{\Lambda}_2\end{bmatrix}\mathbf{C}^{-1}-\mathbf{I}\right]\mathbf{S}^{-1/2}+\mathbf{I}\right| \\
&= \left|\begin{bmatrix}(\mathbf{C}_1^H\mathbf{C}_1)^{-1} & 0\\ 0 & 0\end{bmatrix}\times\begin{bmatrix}\mathbf{\Lambda}_1 & 0\\ 0 & \mathbf{\Lambda}_2\end{bmatrix}-\begin{bmatrix}(\mathbf{C}_1^H\mathbf{C}_1)^{-1} & 0\\ 0 & 0\end{bmatrix}\mathbf{C}^H\mathbf{C}+\mathbf{I}\right| \quad (16)\\
&= \left|\begin{bmatrix}(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1 & 0\\ 0 & 0\end{bmatrix}-\begin{bmatrix}\mathbf{I} & (\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{C}_1^H\mathbf{C}_2\\ 0 & 0\end{bmatrix}+\mathbf{I}\right| \quad (17)\\
&= \left|\begin{bmatrix}(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1 & -(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{C}_1^H\mathbf{C}_2\\ 0 & \mathbf{I}\end{bmatrix}\right| \\
&= \left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1\right| \quad (18)
\end{aligned}$$

where (16) follows from the fact that $|\mathbf{AB}+\mathbf{I}|=|\mathbf{BA}+\mathbf{I}|$, and (17) follows since

$$\mathbf{C}^H\mathbf{C}=[\mathbf{C}_1\quad\mathbf{C}_2]^H[\mathbf{C}_1\quad\mathbf{C}_2]=\begin{bmatrix}\mathbf{C}_1^H\mathbf{C}_1 & \mathbf{C}_1^H\mathbf{C}_2\\ \mathbf{C}_2^H\mathbf{C}_1 & \mathbf{C}_2^H\mathbf{C}_2\end{bmatrix}.$$

We now return to the general optimization problem in (5) with non-zero $\mathbf{K}_w$. As the helper begins to broadcast artificial noise, both the mutual information between the transmitter and the intended receiver $I(X_1;Y_1)$ and the mutual information between the transmitter and the eavesdropper $I(X_1;Y_2)$ are in general decreased. Both of these functions are non-increasing in $\mathbf{K}_w$ since

$$\frac{|\mathbf{A}+\mathbf{B}|}{|\mathbf{B}|}\geq\frac{|\mathbf{A}+\mathbf{B}+\triangle|}{|\mathbf{B}+\triangle|}$$

when $\mathbf{A}$, $\triangle\succeq 0$ and $\mathbf{B}\succ 0$ [20]. A favorable choice for $\mathbf{K}_w$ would be one that reduces $I(X_1;Y_2)$ more than $I(X_1;Y_1)$. Since the optimal solution to (5) is intractable, we propose a suboptimal approach that introduces an additional constraint; namely, we search among those $\mathbf{K}_w$ matrices that guarantee no decrease in the favorable term $I(X_1;Y_1)$ while the power constraint (3) is satisfied. It should be noted that this approach is more general than the cooperative jamming schemes proposed in [10], [11] for the MISO case where the jamming signal is nulled out at the destination. Clearly, such sub-optimal solutions are restricted to the case where there exists a null space between the helper and the intended receiver.

In the following, we obtain an expression that represents all $\mathbf{K}_w\succeq 0$ matrices with the power constraint $\text{Tr}(\mathbf{K}_w)=P_h$ that do not impact the mutual information between the transmitter and the intended receiver; i.e.,

$$I(X_1;Y_1)|_{\mathbf{K}_w\succeq 0,\mathbf{K}_x=\mathbf{K}_x^*}=I(X_1;Y_1)|_{\mathbf{K}_w=0,\mathbf{K}_x=\mathbf{K}_x^*},$$

or from (15)

$$\log\left|\mathbf{K}_x^*\mathbf{H}_1^H(\mathbf{G}_2\mathbf{K}_w\mathbf{G}_2^H+\mathbf{I})^{-1}\mathbf{H}_1+\mathbf{I}\right|=\log\left|\mathbf{K}_x^*\mathbf{H}_1^H\mathbf{H}_1+\mathbf{I}\right|=\log\left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1\right|. \tag{19}$$

Note that, without loss of generality, we have used an equality power constraint $\mathrm{Tr}(\mathbf{K}_w)=P_h$ since for the desired $\mathbf{K}_w$ the best performance is in general obtained when helper transmits at maximum power.

Theorem 1: All $\mathbf{K}_w\succeq 0$ matrices for which $\log\left|\mathbf{K}_x^*\mathbf{H}_1^H(\mathbf{G}_2\mathbf{K}_w\mathbf{G}_2^H+\mathbf{I})^{-1}\mathbf{H}_1+\mathbf{I}\right|=\log\left|\mathbf{K}_x^*\mathbf{H}_1^H\mathbf{H}_1+\mathbf{I}\right|=\log\left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1\right|$ satisfy the following relation:

$$\mathbf{H}_1^H(\mathbf{G}_2\mathbf{K}_w\mathbf{G}_2^H+\mathbf{I})^{-1}\mathbf{H}_1=\mathbf{S}^{-1/2}\left[\mathbf{C}^{-H}\begin{bmatrix}\mathbf{\Lambda}_1 & 0\\ 0 & \mathbf{N}\end{bmatrix}\mathbf{C}^{-1}-\mathbf{I}\right]\mathbf{S}^{-1/2} \tag{20}$$

where

$$\begin{gathered}\mathbf{\Lambda}_{22}\preceq\mathbf{N}\preceq\mathbf{\Lambda}_2\\ \mathbf{\Lambda}_{22}=\mathbf{C}_2^H\mathbf{C}_2+\mathbf{C}_2^H\mathbf{C}_1(\mathbf{\Lambda}_1-\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{C}_1^H\mathbf{C}_2\end{gathered} \tag{21}$$

and $\mathbf{\Lambda}_1$, $\mathbf{\Lambda}_2$, $\mathbf{C}$, $\mathbf{C}_1$ and $\mathbf{C}_2$ are defined in (8)-(11).

Proof: In Appendix A, using similar steps as those used to obtain (18), we show that all $\mathbf{\Sigma}\succeq 0$ matrices for which $\log|\mathbf{K}_x^*\mathbf{\Sigma}+\mathbf{I}|=\log\left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1\right|$ must have the following form

$$\mathbf{\Sigma}=\mathbf{S}^{-1/2}\left[\mathbf{C}^{-H}\begin{bmatrix}\mathbf{\Lambda}_1 & \mathbf{M}\\ \mathbf{M}^H & \mathbf{N}\end{bmatrix}\mathbf{C}^{-1}-\mathbf{I}\right]\mathbf{S}^{-1/2}\,. \tag{22}$$

In the following, we obtain matrices $\mathbf{N}\succeq 0$ and $\mathbf{M}$ and complete the proof by considering the following specific choice for $\mathbf{\Sigma}$:

$$\mathbf{\Sigma}=\mathbf{H}_1^H(\mathbf{G}_2\mathbf{K}_w\mathbf{G}_2^H+\mathbf{I})^{-1}\mathbf{H}_1\,. \tag{23}$$

For the specific $\mathbf{\Sigma}$ in (23), it is evident that

$$0\preceq\mathbf{\Sigma}\preceq\mathbf{H}_1^H\mathbf{H}_1. \tag{24}$$

By applying the constraint $\mathbf{\Sigma}\preceq\mathbf{H}_1^H\mathbf{H}_1$ on (22) and using (14), it is enough to show that:

$$\begin{bmatrix}\mathbf{\Lambda}_1 & \mathbf{M}\\ \mathbf{M}^H & \mathbf{N}\end{bmatrix}\preceq\begin{bmatrix}\mathbf{\Lambda}_1 & 0\\ 0 & \mathbf{\Lambda}_2\end{bmatrix}$$

or equivalently that

$$\begin{bmatrix}0 & -\mathbf{M}\\ -\mathbf{M}^H & \mathbf{\Lambda}_2-\mathbf{N}\end{bmatrix}\succeq 0\,.$$

By applying the Schur Complement Lemma [18], the above relationship is true *iff* $\mathbf{\Lambda}_2 - \mathbf{N} \succeq 0$ and $-\mathbf{M}(\mathbf{\Lambda}_2 - \mathbf{N})^{-1}\mathbf{M}^H \succeq 0$, which in turn is true only when

$$\mathbf{M} = 0 \tag{25}$$

$$\mathbf{\Lambda}_2 - \mathbf{N} \succeq 0\,. \tag{26}$$

Applying the results of (25) and (26) in (22) for the specific choice of $\mathbf{\Sigma} = \mathbf{H}_1^H(\mathbf{G}_2\mathbf{K}_w\mathbf{G}_2^H + \mathbf{I})^{-1}\mathbf{H}_1$, we have:

$$\mathbf{\Sigma} = \mathbf{S}^{-1/2}\left[\mathbf{C}^{-H}\begin{bmatrix}\mathbf{\Lambda}_1 & 0\\ 0 & \mathbf{N}\end{bmatrix}\mathbf{C}^{-1} - \mathbf{I}\right]\mathbf{S}^{-1/2}\,. \tag{27}$$

Based on (24), we also need to show that $\mathbf{\Sigma} \succeq 0$. From (27), it is enough to show that

$$\begin{bmatrix}\mathbf{\Lambda}_1 & 0\\ 0 & \mathbf{N}\end{bmatrix} - \mathbf{C}^H\mathbf{C} = \begin{bmatrix}\mathbf{\Lambda}_1 - \mathbf{C}_1^H\mathbf{C}_1 & -\mathbf{C}_1^H\mathbf{C}_2\\ -\mathbf{C}_2^H\mathbf{C}_1 & \mathbf{N} - \mathbf{C}_2^H\mathbf{C}_2\end{bmatrix} \succeq 0\,.$$

By applying the Schur Complement Lemma, the above relationship is true *iff* $\mathbf{\Lambda}_1 - \mathbf{C}_1^H\mathbf{C}_1 \succeq 0$ and $\mathbf{N} - \mathbf{C}_2^H\mathbf{C}_2 - \mathbf{C}_2^H\mathbf{C}_1(\mathbf{\Lambda}_1 - \mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{C}_1^H\mathbf{C}_2 \succeq 0$. Using Eqs. (8)-(10), it is evident that

$$\mathbf{\Lambda}_1 - \mathbf{C}_1^H\mathbf{C}_1 = \mathbf{C}_1^H\left[\mathbf{S}^{\frac{1}{2}}\mathbf{H}_1^H\mathbf{H}_1\mathbf{S}^{\frac{1}{2}} + \mathbf{I}\right]\mathbf{C}_1 - \mathbf{C}_1^H\mathbf{C}_1 = \mathbf{C}_1^H\mathbf{S}^{\frac{1}{2}}\mathbf{H}_1^H\mathbf{H}_1\mathbf{S}^{\frac{1}{2}}\mathbf{C}_1 \succeq 0$$

and finally the lower bound for $\mathbf{N}$ is given by $\mathbf{N} \succeq \mathbf{C}_2^H\mathbf{C}_2 + \mathbf{C}_2^H\mathbf{C}_1(\mathbf{\Lambda}_1 - \mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{C}_1^H\mathbf{C}_2 \succ 0$, which completes the proof.

It should be noted that as $\mathbf{N} \to \mathbf{\Lambda}_{22}$, we have $\mathrm{Tr}(\mathbf{K}_w) \to \infty$. Moreover, $\mathrm{Tr}(\mathbf{K}_w) = 0$ is achieved by $\mathbf{N} = \mathbf{\Lambda}_2$. Hence, for each scalar P_h, there always exists an $\mathbf{N}$ in the range $\mathbf{\Lambda}_{22} \preceq \mathbf{N} \preceq \mathbf{\Lambda}_2$ that will lead to a $\mathbf{K}_w$ that satisfies (20) with $\mathrm{Tr}(\mathbf{K}_w) = P_h$.

Thus far, we have not made any assumption on the number of antennas at each node. But it is clear from (20) that, for example when $\mathbf{G}_2$ has more columns than rows, for a fixed $\mathbf{N}$ in the acceptable range (21) there will be an infinite number of $\mathbf{K}_w$ matrices that satisfy (20) and consequently do not decrease $I(X_1;Y_1)$. In fact, in this example, a common policy for the helper is to simply transmit artificial noise in the null space of $\mathbf{G}_2$. A more interesting case occurs when no such null space exists, i.e., when the number of antennas at the helper is less than or equal to that of the intended receiver ($n_h \leq n_r$). The above result demonstrates the non-trivial fact that even when $n_h \leq n_r$, it is possible to find a non-zero jamming signal that does not impact $I(X_1;Y_1)$ even when the jamming signal can not be nulled by the channel. In the next section, we find more constructive expressions for the $\mathbf{K}_w$ matrices that satisfy (20) for various combinations of the number of antennas at different nodes. In particular, we show that when $n_h \leq n_r$, a closed-form expression for $\mathbf{K}_w$ can be found.

IV. Results for Different Scenarios

In this section, we consider all possible combinations of the number of antennas at the transmitter, helper and intended receiver, and obtain constructive methods for computing specific $\mathbf{K}_w$ matrices that satisfy (20). Such $\mathbf{K}_w$ will have no impact on $I(X_1;Y_1)$, but will in general decrease $I(X_1;Y_2)$, the mutual information between the transmitter and the eavesdropper, compared with the case that there is no helper. Hence, the secrecy level of the confidential message is increased. As mentioned before, additional improvement in the secrecy rate can be achieved if the transmitter updates its covariance matrix $\mathbf{K}_x$ once $\mathbf{K}_w$ is computed. Note, however, that such an iterative process will not be pursued beyond updating $\mathbf{K}_x$; unlike the first step, where $\mathbf{K}_w$ was updated from its initial value of zero, there is no guarantee that finding a new $\mathbf{K}_w$ will reduce $I(X_1;Y_2)$. Hence, the final secrecy rate for the proposed method is obtained by simply computing (6) and (7) for the resulting $\mathbf{K}_w$ matrices derived in this section.

A. Case 1: $n_h \leq \min\{n_r, n_t\}$

We show here that for the case where $n_h \leq \min\{n_r, n_t\}$ and for a fixed $\mathbf{N}$ in the acceptable range (21), there is only one $\mathbf{K}_w$ matrix that satisfies (20) and consequently does not decrease $I(X_1;Y_1)$. Using the matrix inversion lemma, Eq. (20) can be written as:

$$\begin{aligned}\mathbf{H}_1^H(\mathbf{G}_2\mathbf{K}_w\mathbf{G}_2^H+\mathbf{I})^{-1}\mathbf{H}_1 &= \mathbf{H}_1^H\mathbf{H}_1-\mathbf{H}_1^H\mathbf{G}_2(\mathbf{G}_2^H\mathbf{G}_2+\mathbf{K}_w^{-1})^{-1}\mathbf{G}_2^H\mathbf{H}_1\\ &= \mathbf{S}^{-1/2}\left[\mathbf{C}^{-H}\begin{bmatrix}\mathbf{\Lambda}_1 & 0\\ 0 & \mathbf{N}\end{bmatrix}\mathbf{C}^{-1}-\mathbf{I}\right]\mathbf{S}^{-1/2}.\end{aligned}$$

Replacing $\mathbf{H}_1^H\mathbf{H}_1$ with (14), we have:

$$\mathbf{H}_1^H\mathbf{G}_2(\mathbf{G}_2^H\mathbf{G}_2+\mathbf{K}_w^{-1})^{-1}\mathbf{G}_2^H\mathbf{H}_1=\mathbf{S}^{-1/2}\mathbf{C}^{-H}\begin{bmatrix}0 & 0\\ 0 & \mathbf{\Lambda}_2-\mathbf{N}\end{bmatrix}\mathbf{C}^{-1}\mathbf{S}^{-1/2}. \tag{28}$$

Since we have assumed that the channels are full rank, in the case of $n_h \leq n_r \leq n_t$ or $n_h \leq n_t \leq n_r$, it is clear that rank$(\mathbf{G}_2^H\mathbf{H}_1)=n_h$. Thus, from (28) we have:

$$(\mathbf{G}_2^H\mathbf{G}_2+\mathbf{K}_w^{-1})^{-1}=\mathbf{O}^H\mathbf{S}^{-1/2}\mathbf{C}^{-H}\begin{bmatrix}0 & 0\\ 0 & \mathbf{\Lambda}_2-\mathbf{N}\end{bmatrix}\mathbf{C}^{-1}\mathbf{S}^{-1/2}\mathbf{O} \tag{29}$$

where $\mathbf{O}$ is the right inverse of $\mathbf{G}_2^H\mathbf{H}_1$, which, for example when $n_h \leq n_r \leq n_t$, can be written as $\mathbf{O}=\mathbf{H}_1^H(\mathbf{H}_1\mathbf{H}_1^H)^{-1}\mathbf{G}_2(\mathbf{G}_2^H\mathbf{G}_2)^{-1}$. The following lemma is a direct result of Eqs. (28) and (29).

Lemma 3: For the case of $n_h \leq \min\{n_r, n_t\}$ and for a fixed $\mathbf{N}$ in the acceptable range (21), the $\mathbf{K}_w \succeq 0$ matrix for which (20) is satisfied and $I(X_1;Y_1)$ is not decreased is given by

$$\mathbf{K}_w=\mathbf{Q}-\mathbf{Q}\mathbf{G}_2^H(\mathbf{G}_2\mathbf{Q}\mathbf{G}_2^H-\mathbf{I})^{-1}\mathbf{G}_2\mathbf{Q} \tag{30}$$

where $\mathbf{Q}$ is the RHS of (29).

Proof: After applying the matrix inversion lemma on the LHS of (29), a straightforward computation yields (30).

As is evident from Eqs. (29)-(30), we still have a design parameter, $\mathbf{N}$, that should be chosen in its acceptable range $\mathbf{\Lambda}_{22} \preceq \mathbf{N} \preceq \mathbf{\Lambda}_2$ such that the power constraint $\text{Tr}(\mathbf{K}_w) = P_h$ is satisfied. Finding the optimal $\mathbf{N}$ that minimizes $I(X_1; Y_2)$ when $\mathbf{K}_x$ and $\mathbf{K}_w$ are given by (13) and (30), respectively, is as intractable as the general optimization problem in (5). Instead, we simply restrict the $\mathbf{N}$ we consider to those that can be linearly parameterized within the acceptable range, as follows:

$$\mathbf{N} = \mathbf{\Lambda}_{22} + t\,(\mathbf{\Lambda}_2 - \mathbf{\Lambda}_{22}) \ . \tag{31}$$

Consequently the term $\mathbf{\Lambda}_2 - \mathbf{N}$ in Eq. (30) becomes

$$\mathbf{\Lambda}_2 - \mathbf{N} = (1-t)\,(\mathbf{\Lambda}_2 - \mathbf{\Lambda}_{22})$$

where the scalar $0 \le t \le 1$ is chosen such that the power constraint $\text{Tr}(\mathbf{K}_w) = P_h$ is satisfied. Note that as $t \to 0$ $(\mathbf{N} \to \mathbf{\Lambda}_{22})$ then $\text{Tr}(\mathbf{K}_w) \to \infty$, and as $t \to 1$ $(\mathbf{N} \to \mathbf{\Lambda}_2)$ then $\text{Tr}(\mathbf{K}_w) \to 0$. Thus, we are guaranteed that an acceptable $\mathbf{N}$ can be found in this way.

B. Case 2: $n_h > \min\{n_r, n_t\}$

As mentioned before, for the case of $n_h > n_r$ and for a fixed $\mathbf{N}$ in the acceptable range (21), there are many $\mathbf{K}_w$ matrices that satisfy (20) and consequently do not decrease $I(X_1; Y_1)$. A common policy for the helper in this case is to transmit artificial noise in the null space of $\mathbf{G}_2$. However, as (20) shows, this policy is sufficient but it is not necessary. In other words, it is possible that the optimal $\mathbf{K}_w$ satisfying (20) has elements outside the null space of $\mathbf{G}_2$. Because of the non-linear constraint in (20), finding the optimal $\mathbf{K}_w$ is intractable. A similar discussion applies for the case of $n_t < n_h \le n_r$.

In this section, we present an approach for computing a suitable $\mathbf{K}_w$. Consider the following jamming signal covariance matrix:

$$\mathbf{K}_w = \mathbf{\Gamma}\,\mathbf{\Pi}\,\mathbf{\Gamma}^H \ , \tag{32}$$

where $\mathbf{\Pi}$ is a $d \times d$ positive semidefinite matrix, and $\mathbf{\Gamma}$ is an $n_h \times d$ matrix. For the case of $n_t < n_h \le n_r$ or $n_h > n_r$, we can choose $\mathbf{\Gamma}$ such that $\mathbf{G}_2\,\mathbf{\Gamma}$ is orthogonal to $\mathbf{H}_1\,\mathbf{K}_x^{*\frac{1}{2}}$, i.e., $\mathbf{K}_x^{*\frac{1}{2}}\mathbf{H}_1^H\mathbf{G}_2\,\mathbf{\Gamma} = \mathbb{0}$. For example, $\mathbf{\Gamma}$ can be chosen as the d right singular vectors in the nullspace of $\mathbf{K}_x^{*\frac{1}{2}}\mathbf{H}_1^H\mathbf{G}_2$. Since $\mathbf{K}_x$ will often be rank deficient, the value of d will typically be larger than $n_h - n_t$ for the case of $n_t < n_h \le n_r$, and larger than $n_h - n_r$ for the case of $n_h > n_r$. For this choice of $\mathbf{\Gamma}$, the resulting $\mathbf{K}_w$ in (32) satisfies

(20), and doesn't decrease $I(X_1;Y_1)$ for $\mathbf{N}=\mathbf{\Lambda}_2$, as is clear from (20). Given $\mathbf{\Gamma}$, the choice of $\mathbf{\Pi}$ can be made to maximize the transfer of the "information" in the helper's jamming signal to the eavesdropper. In particular, note that at the eavesdropper, the covariance of the helper's jamming signal will be given by $\mathbf{H}_2\mathbf{\Gamma}\mathbf{\Pi}\mathbf{\Gamma}^H\mathbf{H}_2^H$. If the eigenvalue decomposition of $\mathbf{\Gamma}^H\mathbf{H}_2^H\mathbf{H}_2\mathbf{\Gamma}$ is written as

$$\mathbf{\Gamma}^H\mathbf{H}_2^H\mathbf{H}_2\mathbf{\Gamma}=\mathbf{U}\,\mathbf{D}\,\mathbf{U}^H$$

with $\mathbf{U}$ unitary and $\mathbf{D}$ square and diagonal, then $\mathbf{\Pi}$ can be found via waterfilling; i.e.,

$$\mathbf{\Pi}=\mathbf{U}\,\mathbf{\Delta}\,\mathbf{U}^H\ ,$$

where $\mathbf{\Delta}=\left[\eta\mathbf{I}-\mathbf{D}^{-1}\right]^+$, the operation $[\mathbf{A}]^+$ zeros out any negative elements, and the water-filling level η is chosen such that $\mathrm{Tr}(\mathbf{K}_w)=\mathrm{Tr}(\mathbf{\Delta})=P_h$.

V. Numerical Results

In this section, we present numerical results to illustrate our theoretical findings. In all of the following figures, channels are assumed to be quasi-static flat Rayleigh fading and independent of each other. The channel matrices $\mathbf{H}_1\in\mathbb{C}^{n_r\times n_t}$ and $\mathbf{G}_2\in\mathbb{C}^{n_r\times n_h}$ have i.i.d. entries distributed as $\mathcal{CN}(0,\sigma_d^2)$, while $\mathbf{G}_1\in\mathbb{C}^{n_e\times n_t}$ and $\mathbf{H}_2\in\mathbb{C}^{n_e\times n_h}$ have i.i.d. entries distributed as $\mathcal{CN}(0,\sigma_c^2)$. In each figure, values for the number of antennas at each node, as well as σ_d^2 and σ_c^2, will be depicted. Unless otherwise indicated, results are calculated based on an average of at least 500 independent channel realizations.

In the first example, Fig. 1, we randomly generate positive definite matrices $\mathbf{S}$ such that $\mathrm{Tr}(\mathbf{S})\le P_t$. For each $\mathbf{S}$, we compute the secrecy capacity of the MIMO Gaussian wiretap channel without helper ($\mathbf{K}_w=\mathbb{0}$) as given by (12). Next, using (30), we obtain a $\mathbf{K}_w$ with the average power constraint $\mathrm{Tr}(\mathbf{K}_w)=P_h$ that does not decrease $I(X_1;Y_1)$, and then update $\mathbf{K}_x$ and compute $C_{sec}(\mathbf{S})$, using (6) and (7), accordingly. Fig. 1 compares the secrecy capacity of the wiretap channel with (solid lines) and without (dotted lines) the helper. Note that the vertical difference between the solid curves (about 0.6 bps/channel use) represents the role of the transmit power P_t on the secrecy capacity with helper when P_t changes from 100 to 150 and $P_h=20$. This relatively small difference indicates that, in this example, P_t does not have a big impact on the secrecy capacity. Its role is even more negligible when $P_h=0$, where only an increase of 0.3 bps/channel use is obtained as P_t increases from 100 to 150. The role of the helper on the other hand is significantly more important; increasing P_h from 0 to 20 while holding P_t fixed results in an increase on the order of 3 bps/channel use. Furthermore, the use of the helper with a total power of only 120 ($P_t=100, P_h=20$) provides significantly better secrecy performance than not using the helper and transmitting with total power equal to 150 ($P_t=150, P_h=0$).

In the next examples, we calculate the secrecy capacity of the proposed algorithms under the assumption of an *average* power constraint P_t at the transmitter, and under the constraint that the helper does not reduce the mutual information between the transmitter and receiver. While Eqs. (6) and (7) provide the performance for a specific $\mathbf{S}$, one must solve [17], [20, Lemma 1]

$$C_{sec}(P_t) = \max_{\mathbf{S}\succeq 0, \mathrm{Tr}(\mathbf{S})\leq P_t} C_{sec}(\mathbf{S}) \tag{33}$$

to find the secrecy capacity over all $\mathbf{S}$ that satisfy the average power constraint. In the examples that follow, we perform a numerical search to solve (33) and compute the secrecy capacity.

Fig. 2 shows the secrecy capacity versus P_h for a fixed total average power $P_t + P_h = 110$. In this figure, we consider a situation in which $\sigma_c > \sigma_d$, or in other words where the channel between the transmitter and the intended receiver is weaker than the channel between the transmitter and the eavesdropper, and the channel between the helper and the intended receiver is weaker than the channel between the helper and the eavesdropper. The arrow in the figure shows the secrecy capacity without the helper $(P_h = 0)$. The figure shows that a helper with just a single antenna can provide a dramatic improvement in secrecy rate with very little power allocated to the jamming signal; in fact, the optimal rate is obtained when P_h is less than 2% of the total available transmit power. If the number of antennas at the helper increases, a much higher secrecy rate can be obtained, but at the expense of allocating more power to the helper and less to the signal for the desired user.

In Fig. 3, we consider a situation in which, unlike the above example, we have $\sigma_d > \sigma_c$. Thus, the intended receiver, in comparison with the eavesdropper, receives a weaker information signal and a stronger jamming signal than the eavesdropper. It might seem that in this situation, the helper cannot be very useful, but the figure shows that even in this case we can have a notable improvement in the secrecy rate (about 4 bps/channel use) by increasing the number of antennas at the helper, and with an appropriate power assignment between the transmitter and the helper, without requiring extra total transmit power for the helper node.

In Fig. 4, we consider a specific scenario where the secrecy capacity in the absence of the helper node is zero. While channel matrices $\mathbf{H}_2$ and $\mathbf{G}_2$ are generated randomly with i.i.d. entries distributed as $\mathcal{CN}(0, \sigma_c^2)$ and $\mathcal{CN}(0, \sigma_d^2)$, respectively, we assume the following specific choices for $\mathbf{H}_1$ and $\mathbf{G}_1$:

$$\mathbf{H}_1 = \begin{bmatrix} -0.25 + 0.5i & -0.35 & -1.25 - 0.9i \\ -0.4 + 0.1i & -0.2 + 0.75i & -i \end{bmatrix}$$

$$\mathbf{G}_1 = \begin{bmatrix} 2+0.25i & 1.5+0.5i & 2i \\ 0.25+0.25i & -0.7+1.5i & 0.5+0.33i \\ -1.5 & -0.5-i & -2.9i \end{bmatrix}.$$

Since $\mathbf{H}_1^H\mathbf{H}_1 \preceq \mathbf{G}_1^H\mathbf{G}_1$, all the generalized eigenvalues of the pencil

$$\left(\mathbf{S}^{\frac{1}{2}}\mathbf{H}_1^H\mathbf{H}_1\mathbf{S}^{\frac{1}{2}} + \mathbf{I}\right) - \gamma\left(\mathbf{S}^{\frac{1}{2}}\mathbf{G}_1^H\mathbf{G}_1\mathbf{S}^{\frac{1}{2}} + \mathbf{I}\right)$$

are zero for all $\mathbf{S} \succeq 0$ and consequently, the secrecy capacity without helper will be zero. In this example, we also assume that not only is the total power fixed at $P_t + P_h = 110$, but also the total number of transmit antennas is fixed at $n_t + n_h = 3$. As in the other examples, the secrecy rate of the wiretap channel is considerably improved with the helper. In this case, the best performance is obtained when the helper has only a single antenna.

Finally, in Fig. 5, we consider the role of number of antennas at the helper, n_h, in the secrecy rate for the specific matrix power constraint $\mathbf{S} = \frac{P_t}{n_t}\mathbf{I}$. Note that the solution of Section IV-A applies for $n_h \leq 3$, while the solution of Section IV-B holds for $n_h > 3$. In all cases, we see that the secrecy rate increases considerably as n_h increases.

VI. Conclusions

In this paper, we have studied the Gaussian MIMO Wiretap channel in the presence of an external jammer/helper, where the helper node assists the transmitter by sending artificial noise independent of the source message to confuse the eavesdropper. The jamming signal from the helper is not required to be decoded by the intended receiver and is treated as noise at both the intended receiver and the eavesdropper. We obtained a closed-form relationship for the structure of the helper's artificial noise covariance matrix that guarantees no decrease in the mutual information between the transmitter and the intended receiver. We showed how to find appropriate solutions within this covariance matrix framework that provide very good secrecy rate performance, even when there is no non-trivial null space between the helper and the intended receiver. The proposed scheme is shown to achieve a notable improvement in secrecy rate even for a fixed average total power and a fixed total number of antennas at the transmitter and the helper, without requiring extra power or antennas to be allocated to the helper node.

Appendix A

We are interested in finding a relationship that represents all matrices $\mathbf{\Sigma} \succ 0$ for which

$$\log|\mathbf{K}_x^*\mathbf{\Sigma} + \mathbf{I}| = \log\left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1\right| \ , \tag{34}$$

where

$$\mathbf{K}_x^* = \mathbf{S}^{\frac{1}{2}}\mathbf{C}\begin{bmatrix} (\mathbf{C}_1^H\mathbf{C}_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix}\mathbf{C}^H\mathbf{S}^{\frac{1}{2}} \ . \tag{35}$$

Using the fact that $|\mathbf{AB}+\mathbf{I}| = |\mathbf{BA}+\mathbf{I}|$, it is clear that $\mathbf{\Sigma}$ will have the form $\mathbf{\Sigma} = \mathbf{S}^{-\frac{1}{2}}\mathbf{C}^{-H}\mathbf{X}\mathbf{C}^{-1}\mathbf{S}^{-\frac{1}{2}}$ for some matrix $\mathbf{X} = \mathbf{X}^H$. Substituting this expression for $\mathbf{\Sigma}$ into (34) results in the following equation that must be solved for $\mathbf{X}$:

$$\log\left|\begin{bmatrix} (\mathbf{C}_1^H\mathbf{C}_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix}\mathbf{X}+\mathbf{I}\right| = \log\left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{\Lambda}_1\right| \ . \tag{36}$$

Write $\mathbf{X}$ as $\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \\ \mathbf{X}_2^H & \mathbf{X}_3 \end{bmatrix}$ so that we have

$$\begin{bmatrix} (\mathbf{C}_1^H\mathbf{C}_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix}\mathbf{X}+\mathbf{I} = \begin{bmatrix} (\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{X}_1+\mathbf{I} & (\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{X}_2 \\ 0 & \mathbf{I} \end{bmatrix} ,$$

and note that the determinant of the above matrix is given by $\left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{X}_1+\mathbf{I}\right|$. By comparing this result with (34), we see that $\mathbf{X}_1 = \mathbf{\Lambda}_1 - (\mathbf{C}_1^H\mathbf{C}_1)$. Consequently, we have:

$$\mathbf{\Sigma} = \mathbf{S}^{-\frac{1}{2}}\mathbf{C}^{-H}\begin{bmatrix} \mathbf{\Lambda}_1 - (\mathbf{C}_1^H\mathbf{C}_1) & \mathbf{X}_2 \\ \mathbf{X}_2^H & \mathbf{X}_3 \end{bmatrix}\mathbf{C}^{-1}\mathbf{S}^{-\frac{1}{2}} \tag{37}$$

where $\mathbf{X}_2$ and $\mathbf{X}_3$ are still unknown and must be found as described in the text. It is clear that (37) and (22) are equivalent.

REFERENCES

[1] A. Wyner, "The wire-tap channel," *Bell. Syst. Tech. J.*, vol. 54, no. 8, pp. 1355-1387, Jan. 1975.

[2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, pp. 451-456, Jul. 1978.

[3] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proc. IEEE Int. Symp. Information Theory* Toronto, ON, Canada, Jul. 2008, pp. 524-528.

[4] A. Khisti and G. Wornell, "Secure transmission with multiple antennas II: The MIMOME wiretap channel," to appear, *IEEE Trans. Inf. Theory*, 2010. Available at: http://allegro.mit.edu/pubs/posted/journal/2008-khisti-wornell-it.pdf

[5] S. Ali. A. Fakoorian and A. L. Swindlehurst, "Optimal power allocation for the GSVD based MIMO Gaussian wiretap channel," submitted to *IEEE Trans. Inf. Theory*, Available: http://arxiv.org/abs/1006.1890

[6] T. Liu and S. Shamai (Shitz), "A note on secrecy capacity of the multi-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547-2553, 2009.

[7] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "A MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, Article ID 370970, 8 pages, 2009.

[8] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, pp. 339-348, May 1978.

[9] E. Tekin and A. Yener, "The general Gaussian multiple access and two-way wire-tap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 27352751, Jun. 2008.

[10] L. Dong, Z. Han, A. P. Petropulu, H. V. Poor, "Cooperative jamming for wireless physical layer security", in Proc. of *IEEE Workshop on Statistical Signal Processing*, Cardiff, Wales, U.K. 2009

[11] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Proc.*, vol. 58, NO. 3, pp. 1875-1888, Mar. 2010.

[12] L. Lai and H. El Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 40054019, Sep. 2008.

[13] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "The Gaussian wiretap channel with a helping interferer," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008.

[14] J. Wang and A. Swindlehurst, "Cooperative jamming in MIMO ad hoc networks," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, pp. 1719-1723, Nov., 2009.

[15] E. MolavianJazi, M. Bloch, and J. N. Laneman, "Arbitrary jamming can preclude secure communication," in *Proc. Allerton Conf.* Communications, Control, and Computing, Monticello, IL, Sept. 2009.

[16] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180-2189, June 2008.

[17] Ruoheng Liu, Tie Liu, H. Vincent Poor, and Shlomo Shamai (Shitz), "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, to appear.

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*, University Press, Cambridge, UK, 1985.

[19] S. W. Peters and R. W. Heath, Jr., "Interference alignment via alternating minimization," in Proc. of *IEEE ICASSP*, April 2009, Taiwan.

[20] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936-3964, 2006

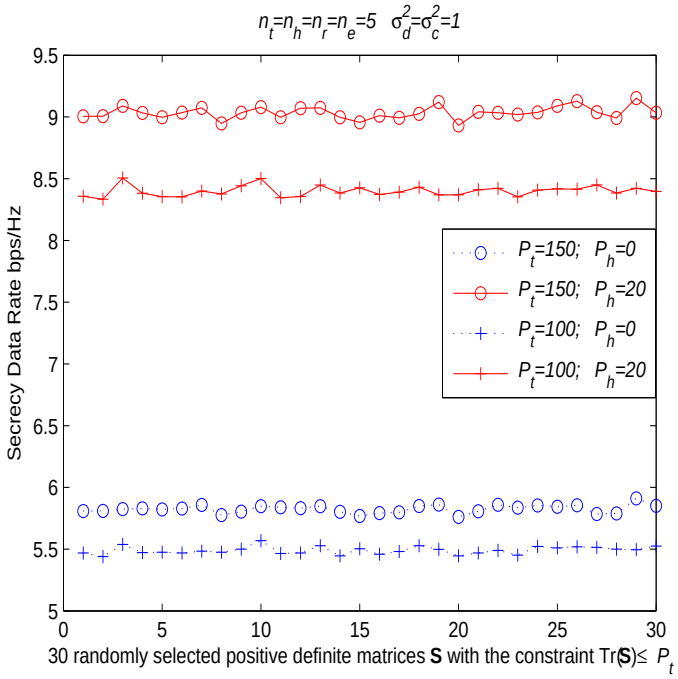

Fig. 1. Comparison of secrecy capacity for MIMO Gaussian wiretap channel with and without helper for different P_t and P_h.

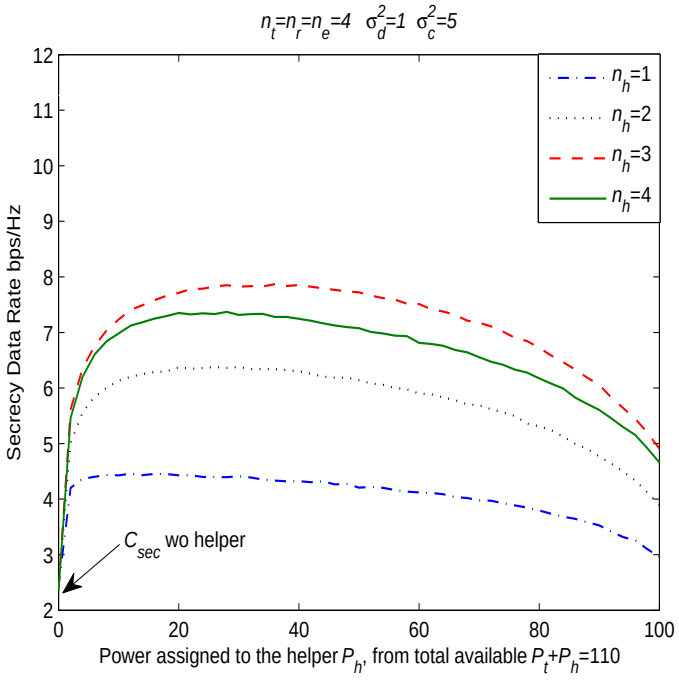

Fig. 2. Comparison of the secrecy capacity for the MIMO Gaussian wiretap channel with and without a helper versus P_h for different number of antennas at the helper, $P_t + P_h = 110$, assuming the eavesdropper's channels are stronger than those of the receiver ($\sigma_d^2 = 1, \sigma_c^2 = 5$).

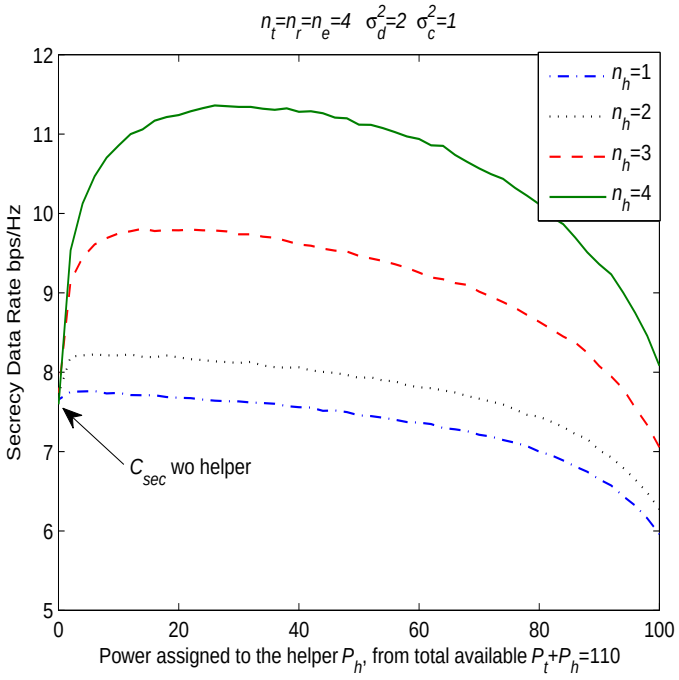

Fig. 3. Comparison of the secrecy capacity for the MIMO Gaussian wiretap channel with and without a helper versus P_h for different number of antennas at the helper, $P_t + P_h = 110$, assuming the receiver's channels are stronger than those of the eavesdropper ($\sigma_d^2 = 2, \sigma_c^2 = 1$).

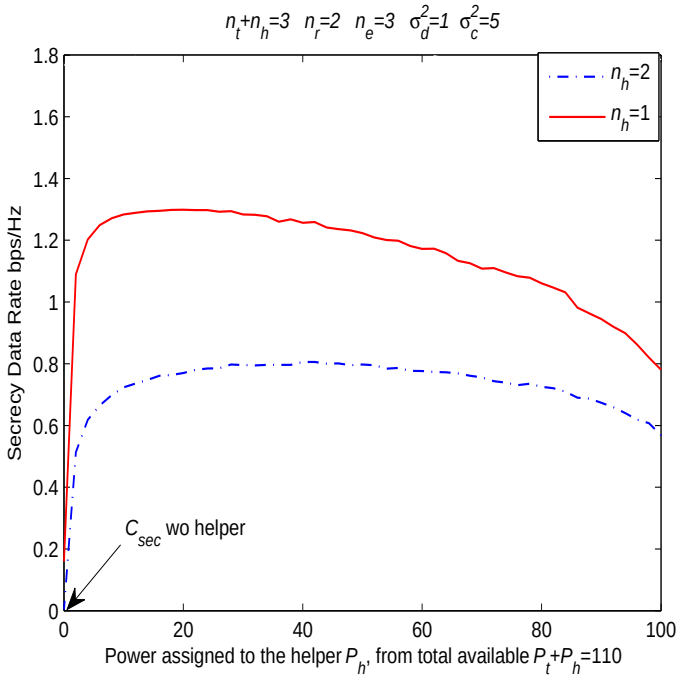

Fig. 4. Comparison of the secrecy capacity for the MIMO Gaussian wiretap channel with and without a helper versus P_h for different number of antennas at the helper, $P_t + P_h = 110$, and $n_t + n_h = 3$.

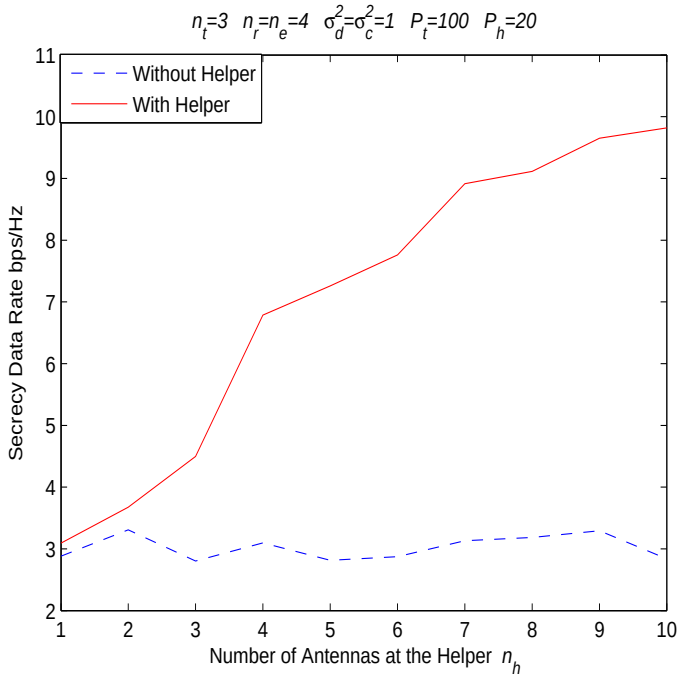

Fig. 5. Secrecy data rate versus n_h for a specific matrix power constraint $\mathbf{S} = \frac{P_t}{n_t}\mathbf{I}$.